Explainable Machine Learning for Predicting Dental Treatment Complexity in Patients with Special Needs and Dentophobia

1st Anita Petreska
Faculty of Information and Communication Technologies - Bitola, University "St. Kliment Ohridski"-
Bitola, Macedonia
e-mail:petreska.anita@uklo.edu.mk

2nd Julijana Nikolovska
Faculty of Dentistry
Skopje, Macedonia
email:
jnikolovska@stomfak.ukim.edu.mk

3rd Efka Zabokova Bilbilova
Faculty of Dentistry
Skopje, Macedonia
email:
ezabokova@stomfak.ukim.edu.mk

4th Vlatko Kokolanski
Faculty of Dentistry
Skopje, Macedonia
email:
vkokolanski@stomfak.ukim.edu.mk

5th Blagoj Ristevski
Faculty of Information and Communication Technologies - Bitola, University "St. Kliment Ohridski"-
Bitola, Macedonia
email: blagoj.ristevski@uklo.edu.mk

6th Nikola Rendevski
Faculty of Information and Communication Technologies - Bitola University "St. Kliment Ohridski"-
Bitola, Macedonia
email: nikola.rendevski@uklo.edu.mk

Abstract—Dental treatment in patients with special needs and severe dentophobia represents a significant clinical and organizational challenge, often requiring individualized planning, including treatment under general anesthesia. This study proposes an explainable machine-learning framework for predicting the complexity of dental treatment using clinical, behavioral, demographic, medical, and socioeconomic data from 303 patients treated at a specialized university dental institution. Treatment complexity was modeled using a predefined binary TARGET variable, with higher-complexity cases representing patients requiring a greater procedural burden according to the clinical database. Logistic Regression, Random Forest, and XGBoost were evaluated using stratified 5-fold cross-validation with ROC-AUC, PR-AUC, and Brier score used to assess discrimination, high-risk case identification, and calibration. Permutation Feature Importance and SHAP analysis were applied to support model interpretability. Logistic Regression was selected as the final model, achieving ROC-AUC = 0.8828 and PR-AUC = 0.9014 with high interpretability. The most influential predictors were primary tooth extractions, permanent tooth extractions, restorative interventions, age, and neurodevelopmental disorders. Subgroup analysis indicated relatively stable performance across sex and financial-status groups, while regional estimates were less stable due to sparse categories. The results suggest that explainable machine learning can support transparent complexity stratification and resource planning in special-care dentistry. However, prospective validation using strictly pre-treatment variables is required before clinical deployment.

Keywords - Explainable Artificial Intelligence, Dental Informatics, Predictive Modeling, Dental Treatment Complexity, Clinical Decision Support Systems.

I. INTRODUCTION

Oral health is an essential component of overall health and quality of life. However, certain patient groups face significant barriers to accessing dental care. Patients with special needs and severe dentophobia represent a particularly vulnerable population, as dental procedures in these individuals are often associated with limited cooperation, high levels of anxiety, and reduced treatment tolerance [21], [22]. As a result, dental therapy is frequently delayed, oral pathology accumulates, and treatment complexity increases. In many cases, these patients require treatment under general anesthesia, creating additional clinical, organizational, and financial challenges [17], [19].

A single factor does not determine dental treatment complexity in this population; rather, it results from the interaction among clinical conditions, behavioral characteristics, socioeconomic circumstances, and prior treatment experiences [22], [23]. In routine clinical practice, complexity assessment is often based on the dentist's experience, which may lead to variability and limited objectivity in treatment planning [19], [21]. Therefore, there is a need for systematic analytical approaches that integrate available data to identify patients at higher risk of complex dental treatment.

Recently, machine learning (ML) has been increasingly applied in medicine and dentistry for prediction, diagnosis, and clinical decision support [1], [5]. In dental research, such approaches have also been used for caries prediction, cavity detection, and treatment-related assessment [7], [8]. However, many models remain difficult to interpret, which limits their practical use in clinical settings where transparency is essential [2], [20]. In this context, explainable ML represents an important step forward, enabling a better understanding of the factors driving model predictions [6], [24].

This study aims to develop an explainable ML framework to identify clinical, behavioral, and sociodemographic factors associated with the complexity of dental treatment among patients with special needs and dentophobia. In addition to predictive modeling, the study includes explainability analysis, fairness assessment, and clinical risk stratification to ensure results are not only statistically valid but also clinically meaningful [3], [25].

The remainder of the paper is structured as follows. The second section surveys the related works, while the subsequent section describes the applied methodology and dataset. The obtained results and the explainability aspects of the proposed approach are described in Section IV. The last section concludes the paper and outlines directions for further research.

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

II. Related Work

Machine learning has seen increasing applications in healthcare, particularly in risk assessment, decision support, and clinical outcome prediction [1], [13]. In dentistry, previous studies have explored artificial intelligence models for endodontic prognosis, caries prediction, oral-health analysis, dental imaging, and restorative dental decision support [4], [5], [7], [8], [12]. Explainable AI methods, including SHAP and permutation-based feature importance, are increasingly relevant because they enhance transparency and trust in clinical prediction models [6], [14]. Broader research has emphasized that explainability is essential not only for model interpretation but also for safe deployment in healthcare environments [2], [3]. However, post hoc explanations should be interpreted cautiously and clinically validated [10], [20]. Despite this progress, relatively few studies have focused on treatment complexity triage in patients with special needs and severe dentophobia. Existing work has mainly examined dental treatment under general anesthesia, retrospective service use, and special-care dentistry guidelines [15], [23]. The present study addresses this gap by integrating predictive modeling, explainability, fairness analysis, and practical clinical triage in a specialized dental care context. Recent work on AI-driven classification of bisphosphonate-related osteonecrosis of the jaw (BRONJ) further supports the relevance of ML-based clinical stratification in oral and maxillofacial conditions [26].

III. Data and Methodology

A. Dataset

The dataset included 303 patients treated in a specialized university dental clinic for individuals with special needs and severe dentophobia. Each record contained 23 predictor variables describing demographic, behavioral, clinical, medical, and socioeconomic characteristics.

The complete predictor set was: Age, Child Gender, Residence Type, Region, Parent Dental Fear, Child Sensory Sensitivity, Previous Negative Dental Experience, Parent Fear Level, Brushing Frequency, Sugar Consumption Frequency, Physical Activity Level, Screen Time, Systemic Disease, Neurodevelopmental Disorders, Allergy, Previous GA, Number Previous GA, Mother Education, Father Education, Family Financial Status, Permanent Restorations, Milk Extractions, and Permanent Extractions. Numerical variables included Age, Child Sensory Sensitivity, Parent Fear Level, Milk Extractions, and Permanent Extractions, while the remaining variables were treated as categorical predictors.

The target variable was dental treatment complexity, represented by the pre-defined binary TARGET field in the clinical dataset. TARGET = 1 denoted higher-complexity cases according to the institutional procedure-burden rule, while TARGET = 0 denoted lower-complexity cases. In the sample, 159 patients (52.48%) were in the higher-complexity group, and 144 (47.52%) were in the lower-complexity group. Because some predictors, such as extraction and restoration counts, are procedure-related, they were interpreted as explanatory and triage-support variables rather than purely pre-treatment causal predictors; this potential target-leakage issue is acknowledged as a limitation and motivates future validation using strictly preoperative variables.

B. Data Preprocessing and Exploratory Analysis

Data quality analysis identified no duplicate records. Missing values were limited, with the highest percentages observed for father's education (6.93%), mother's education (1.98%), and permanent extractions (0.33%), while the remaining variables had little or no missingness. Numerical variables were imputed using the median and standardized. In contrast, categorical variables were converted to strings, imputed using the most frequent category, and one-hot encoded with unknown levels ignored during prediction. For exploratory chi-square analyses, rare categorical levels with fewer than 5 observations were grouped as Other (<5). High-cardinality variables included geographic region (86 categories) and neurodevelopmental diagnostic categories (39 categories). No explicit feature selection or dimensionality reduction was applied before modeling; all 23 predictors were retained, and preprocessing was fitted within each cross-validation fold.

To examine the relationship between predictors and the target variable, several statistical approaches were applied. Numerical variables were analyzed using the Mann–Whitney U test and point-biserial correlation. In contrast, categorical variables were analyzed using the chi-square test of independence and Cramer's V. The strongest statistical association with treatment complexity was observed for primary tooth extractions ($p < 0.001$, Cohen's $d = 0.83$). Age also differed significantly ($p < 0.001$), with patients in the higher-complexity group, on average, younger (15.66 years) than those in the lower-complexity group (19.93 years). Among categorical variables, the strongest association was observed for restorative interventions on permanent teeth ($p < 0.001$, Cramer's $V = 0.387$). Significant associations were also identified with father's education ($p = 0.003$), frequency of sugar consumption ($p = 0.017$), and prior general anesthesia ($p = 0.041$). Point-biserial correlation confirmed that primary tooth extractions had the strongest positive association with higher treatment complexity ($r = 0.384$, $p < 0.001$), while age showed a moderate negative correlation ($r = -0.231$, $p < 0.001$). Pearson correlation analysis also revealed clinically expected relationships: between age and primary tooth extractions ($r = -0.519$) and between age and permanent tooth extractions ($r = 0.302$). In addition, Cramer's V revealed strong associations between prior treatment under general anesthesia and the number of prior procedures under general anesthesia ($V = 0.956$), between type of residence and geographic region ($V = 0.463$), and between systemic diseases and neurodevelopmental disorders ($V = 0.450$).

C. Predictive Modeling

Logistic Regression, Random Forest, and XGBoost were evaluated for predicting dental treatment complexity: These models were selected to compare an interpretable linear approach with more flexible ensemble-based methods capable of modeling nonlinear relationships and interactions [1], [25]. The hyperparameters were fixed before evaluation. Logistic Regression used solver = liblinear, class_weight = balanced, max_iter = 5000, and random_state = 42. Random Forest used n_estimators = 500, min_samples_leaf = 2, class_weight = balanced_subsample, random_state = 42, and n_jobs = -1. XGBoost used n_estimators = 500, learning_rate = 0.05, max_depth = 4, subsample = 0.9, colsample_bytree = 0.9, reg_lambda = 1.0, objective = binary:logistic, eval_metric = logloss, random_state = 42, and n_jobs = -1.

Model performance was assessed using stratified 5-fold cross-validation with shuffling and random_state = 42, preserving class proportions within each fold. Evaluation metrics included ROC-AUC for discrimination, PR-AUC for high-risk case identification, and Brier score for calibration. All preprocessing steps were performed inside the cross-validation pipeline to avoid information leakage. The final

classification threshold (t = 0.3107) was selected from out-of-fold Logistic Regression probabilities by maximizing the F1-score after cross-validation. To assess stability, a nonparametric bootstrap analysis with 2000 repetitions was performed by resampling patient-level out-of-fold predictions with replacement and recalculating the performance metrics.

D. Ethics and Data Availability

The study was based on anonymized and de-identified clinical data obtained from a specialized dental institution. All personal identifiers were removed before analysis. Due to the clinical nature of the data and privacy regulations, the dataset is not publicly available, but aggregated results are presented in this paper.

IV. Results and Explainability

A. Predictive Performance

Among the evaluated models, Logistic Regression and XGBoost demonstrated the best discrimination performance, with nearly identical ROC-AUC values. Logistic Regression achieved an ROC-AUC of 0.8828, while XGBoost achieved 0.8833. Random Forest achieved an ROC-AUC of 0.8453, indicating lower but still acceptable discrimination. The comparative results are presented in Table I.

TABLE I. PREDICTIVE PERFORMANCE OF EVALUATED ML MODELS

Model	ROC-AUC	PR-AUC	Brier Score
Logistic Regression	0.8828	0.9014	0.1400
Random Forest	0.8453	0.8665	0.1869
XGBoost	0.8833	0.8965	0.1395

In terms of PR-AUC, Logistic Regression achieved the highest value (0.9014), followed by XGBoost (0.8965) and Random Forest (0.8665). For calibration, XGBoost achieved the lowest Brier score (0.1395), followed closely by Logistic Regression (0.1400), while Random Forest showed weaker calibration (0.1869). Although XGBoost showed marginally better ROC-AUC and Brier score, Logistic Regression was selected as the final model because it combined strong predictive performance with substantially greater interpretability, which is especially important in clinical decision support. Using the optimized classification threshold t = 0.3107, Logistic Regression achieved a sensitivity of 0.9245, a specificity of 0.6389, and an F1-score of 0.8212. Bootstrap analysis confirmed the model's stability, with a mean ROC-AUC of 0.8822 (95% CI: 0.8447–0.9157), a mean PR-AUC of 0.9008 (95% CI: 0.8637–0.9314), and a mean Brier score of 0.1404 (95% CI: 0.1202–0.1620). Threshold-dependent bootstrap metrics also remained stable, with sensitivity of 0.9243 (95% CI: 0.8817–0.9634), specificity of 0.6373 (95% CI: 0.5577–0.7115), and F1-score of 0.8200 (95% CI: 0.7748–0.8611).

B. Explainability Analysis

To ensure transparency and clinical interpretability, the final model was analyzed using Permutation Feature Importance (PFI) and SHAP [6], [14]. These methods are increasingly used to support clinically understandable ML predictions and structured interpretation of model behavior in healthcare [9], [11]. The PFI analysis showed that the number of primary tooth extractions was the most influential predictor, with the highest mean importance score (0.2421 ± 0.0314). Highly ranked predictors also included permanent tooth extractions, restorative interventions on permanent teeth, age, neurodevelopmental disorders, frequency of sugar consumption, geographic region, and parental education. As shown in Fig. 1, primary tooth extractions, permanent tooth extractions, age, restorative interventions, and neurodevelopmental disorders were among the most influential predictors of higher dental treatment complexity.

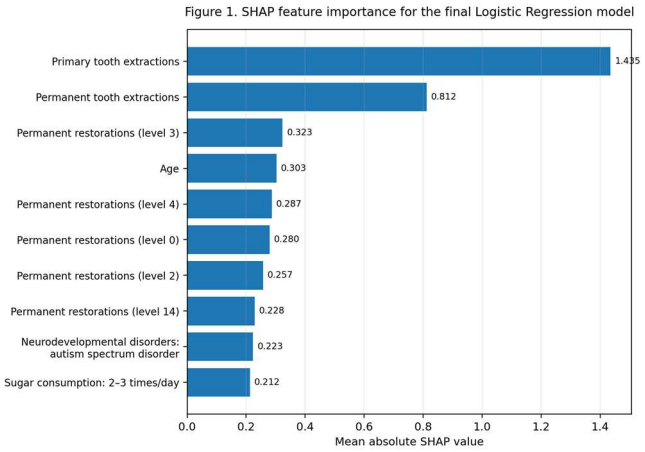


Fig. 1. SHAP feature importance for the final Logistic Regression model.

SHAP analysis confirmed the dominant role of clinical indicators related to accumulated dental pathology. A higher number of primary and permanent tooth extractions increased the probability of higher treatment complexity. Younger age, the presence of neurodevelopmental disorders, restorative interventions, and higher sugar consumption frequency also contributed to increased predicted risk. Additional influences were observed for parental education and prior experience with dental treatment under general anesthesia.

From a clinical perspective, these findings indicate that patients with more extractions and restorative interventions, younger age, and neurodevelopmental disorders are substantially more likely to belong to the higher-complexity group. Overall, the explainability analysis shows that dental treatment complexity is a multifactorial phenomenon arising from the interplay among clinical burden, behavioral habits, and broader socioeconomic context [21], [22].

C. Fairness Analysis

In addition to predictive performance and interpretability, a fairness analysis was conducted to assess whether the model exhibited systematic differences in predictions across demographic and socioeconomic groups. Fairness was assessed using the Demographic Parity Difference (DPD) and the Equal Opportunity Difference (EOD). Subgroups with fewer than 15 patients were excluded from subgroup reporting to reduce instability. Across sex, the model showed DPD = 0.1106 and EOD = 0.0113; male patients (n = 216) and female patients (n = 87) had comparable ROC-AUC values of 0.893 and 0.859, respectively. Across geographic groups, only Skopje (n = 19) and Tetovo (n = 18) met the minimum size criterion, and greater variation was observed (DPD = 0.1988, EOD = 0.2222). These estimates should be interpreted with caution because many regional categories were sparse. Across family financial status, the reported groups were average/managing (n = 194), financial difficulties (n = 77), and comfortable/good (n = 30), with DPD = 0.0557 and EOD = 0.138. Overall, the fairness analysis suggests relatively stable performance across sex and financial groups, while regional differences may reflect local referral pathways, access to care, or sparse subgroup sizes rather than algorithmic bias alone.

D. Clinical Triage

To support practical clinical application, the outputs of the final Logistic Regression model were translated into a simple risk stratification system based on predicted probabilities. The zones were defined as GREEN for predicted risk below 0.33 with low uncertainty, YELLOW for intermediate risk from 0.33 to 0.65 or higher uncertainty, and RED for predicted risk of 0.66 or higher with low uncertainty. The optimized classification threshold t = 0.3107 was used for binary decision support, while the three-zone scheme was used for clinical triage. As shown in Fig. 2, most cases were concentrated in the low-risk (GREEN) and high-risk (RED) regions, while fewer patients were located in the intermediate YELLOW zone. The uncertainty axis appears visually flat because the repeated-estimation uncertainty values were uniformly below 0.08; therefore, the map should primarily be interpreted as a risk-based triage visualization.

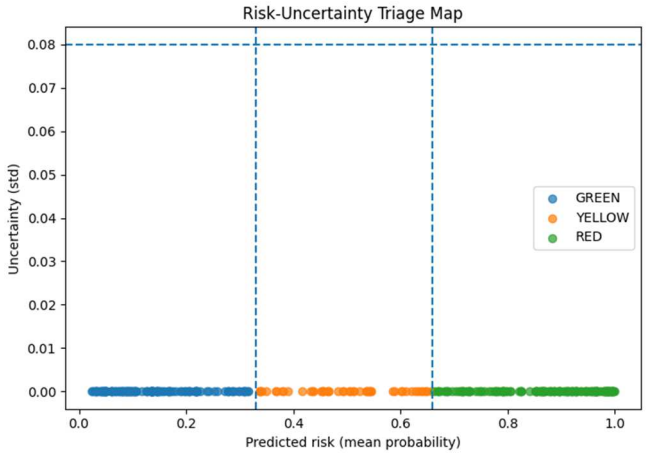


Fig. 2. Risk-based clinical triage map of predicted dental treatment complexity with low estimated uncertainty.

A total of 132 patients were classified in the RED zone, indicating a high risk of complex dental treatment. Forty-nine patients were classified into the YELLOW zone, indicating moderate or uncertain risk, while 122 patients were classified into the GREEN zone, indicating low risk. The probability ranges are adjustable and may be adapted according to clinical preferences, available resources, and the desired trade-off between sensitivity and workload. The triage scheme was not externally validated against independent expert judgment or future clinical outcomes. It should therefore be considered a decision-support tool requiring prospective validation rather than a definitive clinical rule [19], [23].

V. Discussion and Conclusion

The findings of this study show that explainable machine learning can provide meaningful support for understanding and predicting the dental treatment complexity in patients with special needs and severe dentophobia [1], [6]. One key finding is that a relatively simple, highly interpretable model, such as Logistic Regression, can achieve strong discrimination, good calibration, and high sensitivity while remaining clinically understandable [20], [25]. The most influential predictors were indicators of accumulated dental burden, particularly primary tooth extractions, permanent tooth extractions, and restorative interventions. This is clinically plausible, but these variables are also closely related to the outcome definition. Therefore, the results should be interpreted primarily as an explainable triage and resource-planning framework rather than proof of causal pre-treatment prediction. Age and neurodevelopmental disorders also played significant roles, potentially associated with reduced cooperation, sensory hypersensitivity, and greater difficulty during dental procedures [16], [21]. Behavioral and socioeconomic variables, such as sugar consumption frequency, parental education, and geographic region, also contributed to the model's predictions, indicating that treatment complexity is shaped by lifestyle, family context, and access to healthcare [21], [22]. The study is limited by its single-institution design, moderate sample size, sparse regional categories, use of an internally defined target variable, and the absence of external or prospective validation of the triage zones. Future work should validate the model using larger, multicenter datasets, develop models based only on variables available before treatment planning, and compare the triage categories with expert assessments and real-world clinical outcomes.

In conclusion, the results show that explainable ML can provide transparent, clinically interpretable, and practically useful support for treatment planning in vulnerable dental populations [6], [25]. The proposed framework has the potential to contribute to early recognition of complex cases and to a more efficient and equitable organization of dental care for vulnerable populations.

References

[1] Francisco Mesquita, Gonçalo Marques. An explainable machine learning approach for automated medical decision support of heart disease. Data & Knowledge Engineering 153 (2024) 102339

[2] Jeremy Petch, Shuang Di, and Walter Nelson. Opening the Black Box: The Promise and Limitations of Explainable Machine Learning in Cardiology. Canadian Journal of Cardiology 38 (2022) 204e213

[3] Umang Bhatt, Alice Xiang, Shubham Sharma, Adrian Weller, Ankur Taly, Yunhan Jia, Joydeep Ghosh, Ruchir Puri, Jose M. F. Moura, Peter Eckersley. Explainable Machine Learning in Deployment, FAT* '20, January 27–30, 2020, Barcelona, Spain, 647-657.

[4] Yunzi Long, Xiaowei Xu, Jiaqi Chen, Siyi Liu, Jiao Li, Yanmei Dong. An explainable predictive model of direct pulp capping in carious mature permanent teeth. Journal of Dentistry 149 (2024) 105269: 1-7

[5] Mohmed Isaqali Karobari, Abdul Habeeb Adil, Syed Nahid Basheer, Sabari Murugesan, Kamatchi Subramani Savadamoorthi, Mohammed Mustafa, Abdulaziz Abdulwahed, and Ahmed A. Almokhatieb. Evaluation of the Diagnostic and Prognostic Accuracy of Artificial Intelligence in Endodontic Dentistry: A Comprehensive Review of Literature. Hindawi Computational and Mathematical Methods in Medicine Volume 2023, Article ID 7049360, 9 pages

[6] Subrato Bharati, M. Rubaiyat Hossain Mondal, Member, IEEE, and Prajoy Podder. A Review on Explainable Artificial Intelligence for Healthcare: Why, How, and When? Subrato. IEEE Transactions on artificial intelligence, vol. 5, no. 4, april 2024 1429-1442

[7] Priya Dey, Chukwuebuka Ogwo and Marisol Tellez. Comparison of traditional regression modeling vs. AI modeling for predicting dental caries: a secondary data analysis, *Frontiers in Oral Health* 5 (2024): 1322733.

[8] H. Inani, V. Mehta, D. Bhavsar, R. K. Gupta, A. Jain, and Z. Akhtar, "Ai-enabled dental caries detection using transfer learning and gradientbased class activation mapping," Journal of Ambient Intelligence and Humanized Computing, vol. 15, no. 7, pp. 3009–3033, Apr 2024

[9] Eline Stenwig, Giampiero Salvi, Pierluigi Salvo Rossi and Nils Kristian Skjærvold. Comparative analysis of explainable machine learning prediction models for hospital mortality. BMC Medical Research Methodology (2022) 1-14

[10] Emanuele Ratti, Mark Graves. Explainable machine learning practices: opening another black box for reliable medical AI, AI and Ethics (2022) 2:801–814.

[11] Dost Muhammad, Iftikhar Ahmed, Muhammad Ovais Ahmad, and Malika Bendechache Randomized Explainable Machine Learning Models for Efficient Medical Diagnosis, IEEE Journal of Biomedical and Health Informatics, vol. 29, no. 9, September 2025, 6474-6481.

[12] Niharika Bhattacharjee. Automated Dental Cavity Detection System Using Deep Learning and Explainable AI. AMIA Annual Symposium Proceedings. 2022;2022:140–148.

[13] Alvin Har Teck Chia, May Sze Khoo, Andy Zhengyi Lim, Kian Eng Ong, Yixuan Sun, Binh P. Nguyen, Matthew Chin Heng Chua, Junxiong Pang, Explainable machine learning prediction of ICU

mortality. Informatics in Medicine Unlocked Volume 25, 2021, p.100674

[14] Jionghui Lin, Shwetha Hegde, Daochang Liu, Ka Wing Cheng, Heiko Spallek, Chang Xu and Jinlong Gao, Machinometrics, a new framework towards explainable machine learning in healthcare BMC Methods (2025) 2:29.

[15] Shayegan, A., Makanz, V. M., & Abergel, C. (2025). Dental Treatment in Children under General Anesthesia: A 24-year Retrospective Study. International Journal of Clinical Pediatric Dentistry, 18(9), 1134–1139.

[16] Baghdadi, Z. D., Jbara, S., & Muhajarine, N. (2021). Children and parents perspectives on children's dental treatment under general anesthesia: a narratology from Saskatoon, Canada. European Archives of Paediatric Dentistry, 22(4), 725-737.

[17] Puri, S., & Mathew, P. J. (2022). General anesthesia for dental procedures in children: a comprehensive review. Journal of Postgraduate Medicine, Education and Research, 56(1), 29-33.

[18] Schulz-Weidner, N., Schlenz, M. A., Jung, L. G., Uebereck, C. F., Nehls, A., & Krämer, N. (2022). Dental treatment under general anesthesia in pre-school children and schoolchildren with special healthcare needs: a comparative retrospective study. Journal of Clinical Medicine, 11(9), 2613.

[19] Geddis-Regan, A. R., Gray, D., Buckingham, S., Misra, U., Boyle, C., & British Society for Disability and Oral Health. (2022). The use of general anaesthesia in special care dentistry: A clinical guideline from the British Society for Disability and Oral Health. Special Care in Dentistry, 42(S1), 3-32.

[20] Ghassemi, M., Oakden-Rayner, L., & Beam, A. L. (2021). The false hope of current approaches to explainable artificial intelligence in health care. The lancet digital health, 3(11), e745-e750.

[21] Alamri, Hamdan. "Oral care for children with special healthcare needs in dentistry: a literature review." Journal of Clinical Medicine 11, no. 19 (2022): 5557.

[22] Khan, A. J., Sabri, B. A. M., & Ahmad, M. S. (2022). Factors affecting provision of oral health care for people with special health care needs: A systematic review. The Saudi Dental Journal, 34(7), 527-537.

[23] Marinho, Marcello Alves, Flávia Cristina Teixeira Ramos, Andréa Lanzillotti Cardoso, Geraldo Oliveira Silva-Junior, Marcelo Daniel Brito Faria, Luciana Freitas Bastos, Arkadiusz Dziedzic, and Bruna Lavinas Sayed Picciani. "Dental Treatment under General Anesthesia in patients with Special needs provided by private and Public Healthcare Services—A retrospective, comparative study." In Healthcare, vol. 10, no. 6, p. 1147. MDPI, 2022.

[24] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).

[25] Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nat Mach Intell 1: 206–215.

[26] Petreska, A., Ristevski, B., Arsovska, M. M., & Rendevski, N. (2025, July). AI-Driven Classification of Bisphosphonate-Related Osteonecrosis of the Jaw (BRONJ) for Enhanced Clinical Management. In 2025 11th International Conference on Control, Decision and Information Technologies (CoDIT) (Vol. 1, pp. 1324-1330). IEEE.